Stochastic Processes and their Applications 46 (1993) 249-265
North-Holland

Ergodic properties of random measures on stationary sequences of sets

Aaron Gross and James B. Robertson

Department of Statistics and Applied Probability, University of California, Santa Barbara, USA

Received 4 November 1991
Revised 8 April 1992

We study a class of stationary sequences having spectral representation $(M(\tau^n A))_{n\in\mathbb{Z}}$, where A is a set in a measure space $(E, \mathscr{E}, \mu)$, τ is an invertible measure-preserving transformation on $(E, \mathscr{E}, \mu)$, and M is a random measure on $(E, \mathscr{E}, \mu)$. We explore the relationship between the ergodic properties of the sequence and the properties of τ, and construct examples with various ergodic properties using a stacking method on the half-line $[0, \infty)$.

AMS 1990 Subject Classifications: Primary 60G10; Secondary 60G57, 28D05.

mixing * spectral representation * infinitely divisible * stacking method

1. Introduction

In the 1940's and 50's, the ergodic and spectral properties of Gaussian processes were studied extensively, beginning with Maruyama (1949). In 1970, Maruyama provided a spectral representation of infinitely divisible processes with respect to a Poisson measure, and gave a characterization of mixing for infinitely divisible processes in terms of this representation.

More recently, Cambanis, Hardin, and Weron (1987) studied the ergodic properties of symmetric stable processes and Cambanis et al. (1991) studied the ergodic properties of symmetric infinitely divisible processes. Their approach depended on Rajput and Rosinski's (1989) spectral representation of symmetric infinitely divisible processes.

The goal of this paper is to investigate the ergodic properties of a simple class of stochastic processes: namely, stationary sequences which can be represented as the random measure of a stationary sequence of sets. Although the class of sequences with this representation is not large, it does include examples of sequences which are not symmetric infinitely divisible, and hence which were not covered in Cambanis et al. (1991); in the case where the random measure is symmetric infinitely divisible, the main results of Theorem 3.6 are a special case of their results. Example 4.2 is

Correspondence to: Dr. Aaron Gross or Dr. James B. Robertson, Department of Statistics and Applied Probability, University of California, Santa Barbara, CA 93106, USA.

a counter-example to the question posed by Cambanis et al. (1991) of whether weak mixing and mixing are equivalent for non-Gaussian infinitely divisible processes; in fact, it provides a class of such counter-examples which includes symmetric and nonsymmetric α-stable sequences for all $\alpha \in (0, 2]$, Poisson sequences, and many others.

In Section 2 we describe the class of sequences we will study via their spectral representation. In Section 3 we describe the ergodic properties of the sequence in terms of the spectral representation (Theorem 3.6), and we show that ergodicity and weak mixing are equivalent for these sequences, as are mixing and r-mixing of all orders; in addition, we classify those sequences arising from random measures defined on purely atomic measure spaces as either Bernoulli or nonergodic (Theorems 3.1 and 3.5). In Section 4 we use the 'stacking method' (see Friedman, 1970) to construct both mixing sequences and sequences which are weakly mixing but not mixing.

We appreciate several conversations with Aleksander Weron on this subject.

2. Random measures and the spectral representation

The basic set-up is as follows: $(M(\tau^n A))_{n\in\mathbb{Z}}$ is the spectral representation of a stationary sequence, where A is a set of finite measure in a measure space $(E, \mathscr{E}, \mu)$, τ is an invertible measure-preserving transformation on $(E, \mathscr{E}, \mu)$, and M is a random measure on $(E, \mathscr{E}, \mu)$. More specifically, let $(E, \mathscr{E}, \mu)$ be any σ-finite measure space; we do not assume any topological structure on E. We will let $\mathscr{E}_{(f)}$ denote the sets in $\mathscr{E}$ with finite measure. Assume that

(i) M is an independently scattered random (signed) measure on $(E, \mathscr{E}, \mu)$; i.e., M is a real-valued stochastic process $(M(B))_{B\in\mathscr{E}_{(f)}}$ on some probability space $(\Omega, \mathscr{F}, P)$ such that whenever $B_1, B_2, \ldots \in \mathscr{E}_{(f)}$ are disjoint and $\bigcup B_i \in \mathscr{E}_{(f)}$, the random variables $M(B_1)$, $M(B_2), \ldots$ are independent and

$$M(\bigcup B_i) = \sum M(B_i) \quad \text{a.s.},$$

where convergence of the summation may be conditional.

We will also assume that

(ii) M is *stationary*; in that the distribution of $M(B)$ depends only on $\mu(B)$; and

(iii) M is *non-degenerate*; i.e., $M(B)$ is constant only when $\mu(B) = 0$. Note that $M(B) = 0$ whenever $\mu(B) = 0$; consequently we can, and will, interpret set relations in $(E, \mathscr{E}, \mu)$ as holding modulo the null sets.

Let τ be an automorphism (invertible measure-preserving transformation) on $(E, \mathscr{E}, \mu)$. We assume further that

(iv) there is a set $A \in \mathscr{E}_{(f)}$ which *generates* $(E, \mathscr{E})$ under τ, in the sense that $\bigcup \tau^n A = E$ and $\sigma\{\tau^n A\colon n \in \mathbb{Z}\} = \mathscr{E}$. Our goal is to study the sequence $(M(\tau^n A))_{n\in\mathbb{Z}}$. Thus no generality is lost in assuming that A generates $(E, \mathscr{E})$.

We also assume without loss of generality that $(M(B))_{B\in\mathscr{E}_{(f)}}$ is defined on $(\Omega, \mathscr{F}, P)$ where $\Omega=\mathbb{R}^{\mathscr{E}_{(f)}}$ and $\mathscr{F}$ is the σ-field induced by $(M(B))_{B\in\mathscr{E}_{(f)}}$. Since the random measure M was originally motivated by sequences, we will also consider the sub-σ-field

$$\hat{\mathscr{F}}=\sigma\{M(\tau^n A): n\in\mathbb{Z}\}.$$

Note that these σ-fields are not generally equal; for instance, $M(A\cap\tau A)$ is $\mathscr{F}$-measurable but not in general $\hat{\mathscr{F}}$-measurable. Abusing notation, we will still write 'P' for P restricted to $\hat{\mathscr{F}}$.

The transformation τ on $(E, \mathscr{E}, \mu)$ induces a transformation T on $(\Omega, \mathscr{F}, P)$:

$$T\omega(B)=\omega(\tau B)\quad \forall\omega\in\Omega,\ B\in\mathscr{E}_{(f)}.$$

We will write $\hat{T}$ for the restriction of T to $(\Omega, \hat{\mathscr{F}}, P)$; this is the shift transformation on the stationary random sequence $(M(\tau^n A))$. While our main interest is the transformation $\hat{T}$, it will turn out to be convenient to study T in order to study $\hat{T}$.

The transformation T induces the shift operator U_T on $L^2(\Omega, \mathscr{F}, P)$, where

$$U_T f=f\circ T.$$

Similarly, $\hat{T}$ induces $U_{\hat{T}}$ and τ induces U_τ.

If M is an infinitely divisible random measure, then obviously the sequence $(M(\tau^n A))$ will be infinitely divisible, and similarly if M is stable, Gaussian, etc. However, not all such sequences can be represented as above. For instance, observe that if $M(\cdot)$ is centered Gaussian with variance $\mu(\cdot)$, then any sequence with the above representation must be nonnegatively correlated.

Remark. The most natural examples are when μ is continuous (nonatomic), and when μ is purely atomic. In the first case, M will be infinitely divisible, and it is easy to show that M is stochastically continuous, i.e., $M(B_n)$ converges to zero in probability whenever $\mu(B_n)$ converges to zero. In fact, it is not much harder to show, using stationarity and countable additivity, that M is stochastically continuous as long as μ is not purely atomic; the proof is omitted. We will need this fact later in Theorem 3.6.

3. Ergodic properties of the random measure

In this section we give various conditions for a random measure to be ergodic, weakly mixing, mixing, or a K-system. Theorems 3.1 and 3.5 classify M (from an ergodic point of view) for the special case where μ is purely atomic. Theorem 3.6 deals with the more general case. First we present the definitions of various mixing properties, including the case where the measure space is infinite.

An automorphism T on a finite or infinite measure space is *ergodic* if the only T-invariant sets are the empty set and the whole space (modulo null sets).

An automorphism T on a probability space $(\Omega, \mathscr{F}, P)$ is said to be *r-mixing* $(r \geq 1)$ if for any $f_0, f_1, \ldots, f_r \in L^{r+1}(\Omega, \mathscr{F}, P)$,

$$\lim_{n_1,\ldots,n_r\to\infty} \boldsymbol{E}[f_0 \cdot (U_T^{n_1} f_1) \cdots (U_T^{n_1+n_2+\cdots+n_r} f_r)] = \boldsymbol{E}(f_0)\boldsymbol{E}(f_1) \cdots \boldsymbol{E}(f_r). \tag{1}$$

When $r = 1$, this property is simply called *mixing*. The space $L^{r+1}(\Omega, \mathscr{F}, P)$ in the above definition can be replaced by any collection of functions which generates it, in particular the indicator functions.

An automorphism T on a probability space is said to be *weakly mixing* if for any L^2 functions f and g,

$$\langle f, U_T^n g\rangle \to \langle f, 1\rangle\langle 1, g\rangle \tag{2}$$

as $n \to \infty$ outside a set of density zero. (A subset J of the natural numbers has *density zero* if

$$\lim_{n\to\infty} \frac{1}{n} \operatorname{card}\{j \in J\colon j \leq n\} = 0.$$

If

$$\limsup_{n\to\infty} \frac{1}{n} \operatorname{card}\{j \in J\colon j \leq n\} > 0,$$

then J is said to have *positive upper density*.)

In the rest of this section, we will examine the relationship between the ergodic properties of T and τ. The automorphism τ, however, may be defined on an infinite (σ-finite) measure space. Krengel and Sucheston (1969) defined mixing for an arbitrary σ-finite measure space; in the infinite case, their definition is equivalent to (1) for $r = 1$, but with the right hand side replaced by zero. Analogously, we define weak mixing for an *infinite* (σ-finite) measure space by (2), but with the right hand side replaced by zero. (Obviously, if one were to define r-mixing by analogy with (1), then any mixing transformation on an infinite measure space would be r-mixing.)

Remark. An automorphism τ on an infinite σ-finite measure space is weakly mixing if and only if there is no τ-invariant set of positive finite measure. In fact, an automorphism τ on a σ-finite measure space is weakly mixing if and only if U_τ has continuous spectrum — the proof is the same as in the case of a probability space (see, for instance, Walters, 1982). If U_τ does not have continuous spectrum, then there is an L^2 eigenfunction f, so $|f|$ is invariant but not constant. Thus for some Borel set C, $|f|^{-1}(C)$ is the desired set. Conversely, if B is an invariant set of finite positive measure, take $f = \mathbf{1}_B$ to see that τ is not weakly mixing. (The preceding argument also shows that on an infinite measure space, ergodicity implies weak mixing.)

An automorphism T on a probability space $(\Omega, \mathscr{F}, P)$ is a *K-automorphism* ('K' for 'Kolmogorov') if there is a σ-field $\mathscr{G} \subset \mathscr{F}$ such that

(i) $T\mathscr{G} \supset \mathscr{G}$,

(ii) $\bigvee T^n\mathscr{G} = \mathscr{F}$,

(iii) $\bigwedge T^n\mathscr{G} = \{\emptyset, \Omega\}$.

An automorphism T on a probability space $(\Omega, \mathscr{F}, P)$ is said to be a *Bernoulli automorphism* if there is a σ-field $\mathscr{G}$ such that the σ-fields $T^n\mathscr{G}$ are independent and generate $\mathscr{F}$.

The following two theorems show that if μ is purely atomic, then T and $\hat{T}$ are either Bernoulli automorphisms, or they are not ergodic.

Theorem 3.1. *Assume that $(E, \mathscr{E}, \mu)$, M, A and τ satisfy assumptions* (i)–(iv) *of Section* 2 *and that T and $\hat{T}$ are induced by τ as described in Section* 2.

If there is an atom contained in an infinite number of $\tau^n A$'s then $\hat{T}$ is not ergodic (and so neither is T).

Proof. By Lemma 3.2, there is an atom D with $\tau^k D = D$ for some $k \neq 0$. Now for each n which is a multiple of k, apply Lemma 3.3 with

$$U = M(A \backslash \tau^n A), \quad V = M(\tau^n A \backslash A), \quad W = M(D), \quad R = M(A \cap \tau^n A \backslash D),$$

to get

$$P(M(A) \leqslant c, M(\tau^n A) \leqslant c) - P(M(A) \leqslant c)^2 \geqslant \delta > 0,$$

where c and δ do not depend on n. By Lemma 3.4, $\hat{T}$ is not ergodic. $\square$

Lemma 3.2. *Assume that $(E, \mathscr{E}, \mu)$ and τ satisfy assumption* (iv) *of Section* 2.

If there is an atom D contained in an infinite number of $\tau^n A$'s, then $\tau^k D = D$ for some $k \neq 0$.

Proof. If $\tau^n D \subset A$ for infinitely many n, then the $\tau^n D$'s could not be disjoint — if they were, then since they each have the same positive measure, A would have infinite measure. This implies that $\tau^k D$ intersects D for some nonzero k. But D is an atom, so $\tau^k D = D$. $\square$

Lemma 3.3. *Let U, V, W, and R be independent random variables with U, V, and W nonconstant and with U and V having the same distribution. Then there are real numbers c and δ, depending only on the distributions of $U + R$ and W, such that*

$$P(U + W + R \leqslant c, V + W + R \leqslant c) - P(U + W + R \leqslant c)^2 \geqslant \delta > 0.$$

Proof. Let R' and R'' have the same distribution as R, and let them be independent of U, V, W, and R. We will choose c so that

$$P(U+W+R\leq c,\ V+W+R\leq c)\geq P(U+W+R'\leq c,\ V+W+R''\leq c)$$
$$> P(U+W+R\leq c)^2. \tag{3}$$

By independence and equality of distributions, for any c,

$$\begin{aligned}&P(U+W+R\leq c,\ V+W+R\leq c)\\&=\int P(W\in \mathrm{d}w)P(U+R\leq c-w,\ V+R\leq c-w\mid W=w)\\&=\int P(W\in\mathrm{d}w)\int P(R\in\mathrm{d}r)P(U\leq c-w-r)^2.\end{aligned}$$

But

$$\begin{aligned}&P(U+W+R'\leq c,\ V+W+R''\leq c)\\&=\int P(W\in\mathrm{d}w)P(U+R'\leq c-w)P(V+R''\leq c-w)\\&=\int P(W\in\mathrm{d}w)P(U+R\leq c-w)^2\\&=\int P(W\in\mathrm{d}w)\left[\int P(R\in\mathrm{d}r)P(U\leq c-w-r)\right]^2.\end{aligned} \tag{4}$$

Therefore the first inequality in (3) follows from Jensen's inequality. Furthermore,

$$P(U+W+R\leq c)^2=\left[\int P(W\in\mathrm{d}w)P(U+R\leq c-w)\right]^2.$$

By Jensen's inequality, this is less than or equal to expression (4). Therefore, in order to get strict inequality in (3) we must find c such that $P(U+R\leq c-w)$ is not almost surely constant as a function of w. Now, there is a real number a such that $P(U+R\leq\cdot)$ is not constant in any neighborhood of a. Since W is not constant, we can choose b such that $P(W<b)$ and $P(W>b)$ are both positive. Now let $c=a+b$; then $P(U+R\leq c-w)$ will take different values for w less than or greater than b and (3) holds. Finally, take

$$\delta=P(U+W+R'\leq c,\ V+W+R''\leq c)-P(U+W+R\leq c)^2. \qquad \square$$

Lemma 3.4. *Assume that* $(E,\mathscr{E},\mu)$, M, A *and* τ *satisfy assumptions* (i)–(iv) *of Section* 2 *and that* T *and* $\hat{T}$ *are induced by* τ *as described in Section* 2.

If there exist real numbers $\delta > 0$ and c such that for all n in a set $J \subset \mathbb{N}$ of positive upper density,

$$P(M(A) \leqslant c, M(\tau^n A) \leqslant c) - P(M(A) \leqslant c)^2 \geqslant \delta,$$

then $\hat{T}$ is not ergodic (and hence neither is T).

Proof. By well-known properties of sets of density zero (Walters, 1982), it follows from the hypothesis that

$$\frac{1}{n} \sum_{k=0}^{n-1} |P(M(A) \leqslant c, M(\tau^n A) \leqslant c) - P(M(A) \leqslant c)^2| \nrightarrow 0.$$

But by Lemma 3.3, with

$$U = M(A \backslash \tau^n A), \quad V = M(\tau^n A \backslash A), \quad W = M(A \cap \tau^n A), \quad R = 0,$$

the expression inside the absolute value sign is always nonnegative. Therefore,

$$\frac{1}{n} \sum_{k=0}^{n-1} P(M(A) \leqslant c, M(\tau^n A) \leqslant c) - P(M(A) \leqslant c)^2 \nrightarrow 0,$$

and this implies $\hat{T}$ is not ergodic. □

Theorem 3.5. *Assume that $(E, \mathscr{E}, \mu)$, M, A and τ satisfy assumptions* (i)–(iv) *of Section* 2 *and that T and $\hat{T}$ are induced by τ as described in Section* 2.

Suppose μ is purely atomic. If every atom is contained in only finitely many $\tau^n A$'s, then T is a Bernoulli automorphism (and hence so is $\hat{T}$).

Proof. Let $\mathscr{G}$ be the σ-field generated by the set of all random variables of the form $M(D)$, where D is an atom contained in A but not in any $\tau^m A$ for $m < 0$. Then the σ-fields $T^n \mathscr{G}$ are independent and generate $\mathscr{F}$, so T is a Bernoulli automorphism. It is well known that a Bernoulli automorphism restricted to a sub-σ-field is also Bernoulli (Walters, 1982). □

We now proceed to the case where $(E, \mathscr{E}, \mu)$ is an arbitrary measure space.

For any real-valued process $(h_\gamma)_{\gamma \in \Gamma}$ on a given measure space, $L^2(h_\gamma)_{\gamma \in \Gamma}$ will denote the space of all complex-valued L^2 functions defined on the measure space and measurable with respect to the σ-field induced by $(h_\gamma)_{\gamma \in \Gamma}$. For the proof of the following theorem, we will need the fact that functions of the form

$$f = \exp\left(i \sum_{l \in L} b_l h_l \right), \tag{5}$$

where $L \subset \Gamma$ is finite and b_l is a real number for each $l \in L$, generate $L^2(h_\gamma)_{\gamma \in \Gamma}$. This follows from the Stone–Weierstrass Theorem and standard approximation arguments.

For convenience, we will use the following abbreviations in the next theorem.

ERG($\cdot$):	The automorphism ($\cdot$) is ergodic.
WMIX($\cdot$):	The automorphism ($\cdot$) is weakly mixing.
WMIX(τ, A):	The sequences $\mu(A \cap \tau^n A)$ converges to zero as n approaches infinity outside some set of density zero.
MIX($\cdot$):	The automorphism ($\cdot$) is mixing.
MIX(τ, A):	The sequence $\mu(A \cap \tau^n A)$ converges to zero.
RMIX($\cdot$):	The automorphism ($\cdot$) is r-mixing for all $r \geq 1$.
K($\cdot$):	The automorphism ($\cdot$) is a K-automorphism.

Theorem 3.6. *Assume that* $(E, \mathscr{E}, \mu)$, M, A *and* τ *satisfy assumptions* (i)–(iv) *of Section* 2 *and that* T *and* $\hat{T}$ *are induced by* τ *as described in Section* 2.

If $0 < \mu(E) < \infty$, *then* $\hat{T}$ *is not ergodic (and hence neither is* T*); if* $\mu(E) = \infty$, *then the following implications hold.*

$$
\begin{array}{ccccccc}
 & & \mathrm{ERG}(\tau) & \Rightarrow & \mathrm{ERG}(T) & \Leftrightarrow & \mathrm{ERG}(\hat{T}) \\
 & & \Downarrow & & \Updownarrow & & \Updownarrow \\
\mathrm{WMIX}(\tau, A) & \Leftrightarrow & \mathrm{WMIX}(\tau) & \Leftrightarrow & \mathrm{WMIX}(T) & \Leftrightarrow & \mathrm{WMIX}(\hat{T}) \\
\Uparrow & & \Uparrow & & \Uparrow & & \Uparrow \\
\mathrm{MIX}(\tau, A) & \Leftrightarrow & \mathrm{MIX}(\tau) & \Leftrightarrow & \mathrm{MIX}(T) & \Leftrightarrow & \mathrm{MIX}(\hat{T}) \\
 & & & & \Updownarrow & & \Updownarrow \\
 & & \Uparrow & & \mathrm{RMIX}(T) & \Leftrightarrow & \mathrm{RMIX}(\hat{T}) \\
 & & & & \Uparrow & & \Uparrow \\
 & & \mu(\limsup \tau^n A) = 0 & \Rightarrow & \mathrm{K}(T) & \Rightarrow & \mathrm{K}(\hat{T})
\end{array}
$$

Remark. The more significant results shown in the above diagram are as follows: the ergodic properties of a sequence are determined by the behavior of $\mu(A \cap \tau^n A)$; all ergodic sequences with the spectral representation $(M(\tau^n A))$ as described above are weakly mixing; all mixing sequences with this spectral representation are r-mixing for all $r \geq 1$, if $\hat{T}$ is mixing, weakly mixing, etc., then so is T.

In the case of the implications involving the K-property, we do not know whether the converses hold. In all other cases, the converses of the single-arrow implications do not hold. Example 4.1 is a counter-example for the converse implications regarding ergodicity of τ; Example 4.2 shows that weak mixing does not imply mixing; and the remark following Example 4.3 shows that τ can be mixing although $\limsup \tau^n A$ has nonzero measure.

Proof of Theorem 3.6. The following implications require no further proof: ERG(τ) implies WMIX(τ) on an infinite measure space (see remarks following the definitions of weak mixing); and each condition in the third column implies the condition above it and the condition to its right (well-known).

We will break the proof into several steps, beginning with the statement concerning the case where E has finite measure.

Step 1: If $0 < \mu(E) < \infty$, then $\hat{T}$ is not ergodic.

Apply Lemma 3.7 with $C_n = \tau^n A$ to get an $\varepsilon > 0$ and a set J of positive upper density such that

$$\mu(A \cap \tau^n A) > \varepsilon \quad \forall n \in J.$$

Then by Lemma 3.8, there exist $c \in \mathbb{R}$ and $\delta > 0$ such that

$$P(M(A) \leqslant c, M(\tau^n A) \leqslant c) - P(M(A) \leqslant c)^2 \geqslant \delta$$

for all n in a set of positive upper density. By Lemma 3.4, $\hat{T}$ is not ergodic.

Step 2:

$$\mu(\limsup \tau^n A) = 0 \quad \Rightarrow \quad \mathrm{K}(T) \quad \Rightarrow \quad \mathrm{K}(\hat{T})$$

If μ is purely atomic, then the hypothesis implies that every atom is contained in only finitely many $\tau^n A$'s, and Theorem 3.5 implies that T is a Bernoulli automorphism (and hence a K-automorphism).

Suppose now that μ is not purely atomic; then by the remarks in Section 2, M is stochastically continuous.

Let $\mathscr{E}^n = \sigma\{\tau^k A: k \geqslant n\}$, and let $\mathscr{F}^n = \sigma\{M(B): B \in \mathscr{E}^n_{(\mathrm{f})}\}$. Clearly $\mathscr{F}^0$ satisfies conditions (i) and (ii) in the definitions of a K-automorphism. Write $\mathscr{F}^\infty = \bigwedge T^n \mathscr{F}^0 = \bigwedge \mathscr{F}^n$. We must show $\mathscr{F}^\infty$ is trivial.

We first establish the following: for any $B \in \mathscr{E}_{(\mathrm{f})}$, $\mu(B \cap C_n)$ converges to zero for any sequence (C_n) with $C_n \in \mathscr{E}^n_{(\mathrm{f})}$. Indeed, B can be approximated by a finite disjoint union of cylinder sets (each contained in some $\tau^j A$), and each C_n can be approximated similarly with each cylinder set contained in some $\tau^j A$, $j \geqslant n$. Thus our claim will be true if $\mu(\bigcup_{j \geqslant n}(A \cap \tau^j A))$ converges to zero. But this follows from the fact that $\limsup \tau^n A$ has measure zero.

Now let f be any element of $L^2(\mathscr{F}^\infty)$, and let g be any element of $L^2(\mathscr{F})$ of the form

$$g = \exp\left(\mathrm{i} \sum_{l \in L} b_l M(B_l)\right). \tag{6}$$

Since $f \in L^2(\mathscr{F}^n)$ for each integer n, there is a sequence $(f_n)_{n \geqslant 1}$ such that f_n converges to f in L^2 and f_n depends only on a finite collection $(M(B_k))_{k \in K_n}$ where $\{B_k: k \in K_n\}$ is a finite subset of $\mathscr{E}^n_{(\mathrm{f})}$. Write

$$C_n = \bigcup_{k \in K_n} B_k.$$

Now define

$$g_n = \exp\left(\mathrm{i} \sum_{l \in L} b_l M(B_l \setminus C_n)\right).$$

Then by independence,

$$\langle f_n, g_n \rangle = \langle f_n, 1 \rangle \langle 1, g_n \rangle \quad \forall n \geqslant 1.$$

But by the observation above, $\mu(B_l \cap C_n)$ converges to zero; this together with stochastic continuity implies that for all $l \in L$,

$$M(B_l \backslash C_n) \to M(B_l) \quad \text{in probability as } n \to \infty,$$

and by dominated convergence, g_n converges to g in L^2; thus

$$\langle f, g\rangle = \langle f, 1\rangle\langle 1, g\rangle.$$

Functions of the form in (6) generate $L^2(\mathscr{F})$, so f is constant and $\mathscr{F}^\infty$ is trivial.

Step 3:

$$\begin{array}{ccccccc} \mathrm{MIX}(\tau, A) & \Leftrightarrow & \mathrm{MIX}(\tau) & \Leftrightarrow & \mathrm{MIX}(T) & \Leftrightarrow & \mathrm{MIX}(\hat{T}) \\ & & & & \Updownarrow & & \Updownarrow \\ & & & & \mathrm{RMIX}(T) & \Leftrightarrow & \mathrm{RMIX}(\hat{T}) \end{array}$$

Step 3(a): $\mathrm{RMIX}(T) \Rightarrow \mathrm{MIX}(\hat{T})$. Trivially true, since r-mixing always implies mixing and since $\hat{\mathscr{F}} \subset \mathscr{F}$.

Step 3(b): $\mathrm{MIX}(\hat{T}) \Rightarrow \mathrm{MIX}(\tau, A)$. Suppose $\mu(A \cap \tau^n A)$ does not converge to zero. Then by Lemma 3.8, there is a real number c such that

$$P(M(A) \leq c, M(\tau^n A) \leq c) \not\to P(M(A) \leq c)^2,$$

and thus $\hat{T}$ is not mixing.

Step 3(c): $\mathrm{MIX}(\tau, A) \Rightarrow \mathrm{MIX}(\tau)$. If $\mu(A \cap \tau^n A)$ goes to zero, then $\mu(F_1 \cap \tau^n F_2)$ converges to zero whenever F_1 and F_2 are cylinder sets, each contained in at least one $\tau^n A$. But these cylinder sets generate $\mathscr{E}$, so by a standard approximation argument (Walters, 1982), τ is mixing.

Step 3(d): $\mathrm{MIX}(\tau) \Rightarrow \mathrm{RMIX}(T)$. If μ is purely atomic and τ is mixing, then every atom is contained in only finitely many $\tau^n A$'s, and by Theorem 3.5 T is Bernoulli and hence r-mixing for all $r \geq 1$.

Suppose now that μ is not purely atomic; then by the remarks in Section 2, M is stochastically continuous. It suffices to prove (1) for a collection of functions which generates L^{r+1}, so let $f_0, f_1, \ldots, f_r$ be functions of the form

$$f_m = \exp\left(\mathrm{i} \sum_{l \in L_m} b_l M(B_{l,m})\right),$$

where for each $m = 0, 1, \ldots, r$, f_m depends only on $(M(B_{l,m}))_{l \in L_m}$, L_m finite (cf. Cambanis et al.'s (1991) 'dynamical functional'). We will approximate the $U_T^{n_1+\cdots+n_m} f_m$'s in (1) by functions which are independent of each other.

Let $n = (n_1, \ldots, n_r)$ in (1). Define

$$C_m^{(n)} = \left(\bigcup_{i=0}^{m-1} \bigcup_{l \in L_i} \tau^{n_1+\cdots+n_i} B_{l,i}\right).$$

$U_T^{n_1+\cdots+n_m} f_m$ is a function of $(M(\tau^{n_1+\cdots+n_m} B_l))_{l \in L_m}$, so define $f_m^{(n)}$ to be $U_T^{n_1+\cdots+n_m} f_m$, but with $M(\tau^{n_1+\cdots+n_m} B_l)$ replaced by

$$M(\tau^{n_1+\cdots+n_m} B \backslash C_m^{(n)}).$$

Since τ is mixing,

$$\mu(\tau^{n_1+\cdots+n_m}B\cap C_m^{(n)})\to 0 \quad \text{as } n_m\to\infty.$$

(The expression above does not depend on $n_1,\ldots,n_{m-1}$.)

M is stochastically continuous, so this implies that for $m=1,\ldots,r$,

$$U_T^{n_1+\cdots+n_m}f_m-f_m^{(n)}\to 0 \quad \text{in probability as } n_m\to\infty,$$

and convergence holds in L^{r+1} as well (dominated convergence). A standard triangle inequality argument then shows that

$$E|f_0U_T^{n_1}f_1\cdots U_T^{n_1+\cdots+n_r}f_r-f_0f_1^{(n)}\cdots f_r^{(n)}|\to 0 \quad \text{as } n_1,\ldots,n_r\to\infty.$$

But the $f_m^{(n)}$'s, $m=0,1,\ldots,r$, are independent, so

$$E(f_0f_1^{(n)}\cdots f_r^{(n)})=Ef_0^{(n)}\cdots Ef_r^{(n)}\to Ef_0\cdots Ef_r$$

as $n_1,\ldots,n_r\to\infty$. Therefore (1) holds and T is r-mixing.

The implications in (a)-(d), along with the implications stated in the remarks at the beginning of the proof of the theorem, establish the assertion of Step 3.

Step 4:

$$\begin{array}{ccccccc}
 & & \mathrm{ERG}(\tau) & \Rightarrow & \mathrm{ERG}(T) & \Leftrightarrow & \mathrm{ERG}(\hat{T}) \\
 & & \Downarrow & & \Updownarrow & & \Updownarrow \\
\mathrm{WMIX}(\tau,A) & \Leftrightarrow & \mathrm{WMIX}(\tau) & \Leftrightarrow & \mathrm{WMIX}(T) & \Leftrightarrow & \mathrm{WMIX}(\hat{T})
\end{array}$$

Step 4(a): WMIX$(\tau,A)\Rightarrow$WMIX(τ). The proof is similar to the proof that MIX(τ,A) implies MIX(τ), except that convergence is for n outside a set of density zero; in order to verify convergence outside a set of density zero for the cylinder sets, observe that a finite union of sets of density zero has density zero.

Step 4(b): WMIX$(\tau)\Rightarrow$WMIX(T). Again, similar to the proof in Step 3.

Step 4(c): WMIX$(T)\Rightarrow$WMIX$(\hat{T})$. This is trivially true, since $\hat{\mathscr{F}}\subset\mathscr{F}$.

Step 4(d): ERG$(\hat{T})\Rightarrow$WMIX(τ,A). Suppose $\mu(A\cap\tau^nA)$ does not converge to zero as n goes to infinity outside any set of density zero. Then τ is not weakly mixing, so by the remarks following the definition of weak mixing there is a τ-invariant subset B of positive finite measure. Apply Lemma 3.7 with $C_n=B\cap\tau^nA$, observing that $\mu(C_n)$ does not depend on n and that $B\subset\bigcup C_n$. Then there is an $\varepsilon>0$ and a set J of positive upper density such that

$$\mu(A\cap\tau^nA)\geqslant\mu(C_l\cap C_n)>\varepsilon \quad \forall n\in J.$$

Now by Lemma 3.8, there exist $\delta>0$ and c such that

$$P(M(A)\leqslant c,\ M(\tau^nA)\leqslant c)-P(M(A)\leqslant c)^2\geqslant\delta$$

for all but finitely many n in J. Hence by Lemma 3.4, $\hat{T}$ is not ergodic.

The assertion of Step 4 follows from (a)-(d) and from the remarks at the beginning of the proof of the theorem. $\square$

Lemma 3.7. *If $(C_n)_{n\in\mathbb{Z}}$ is a sequence of sets in a measure space $(E, \mathscr{E}, \mu)$, each having the same positive finite measure, and $\bigcup C_n$ also has finite measure, then there is an $\varepsilon > 0$, $l \in \mathbb{Z}$, and a set $J \subset \mathbb{N}$ of positive upper density such that*

$$\mu(C_l \cap C_n) > \varepsilon \quad \forall n \in J.$$

Proof. There is an $N \geq 0$ such that

$$\mu\left[\left(\bigcup_{l\in\mathbb{Z}} C_l\right) \Big\backslash \left(\bigcup_{l=-N}^{N} C_l\right)\right] < \tfrac{1}{2}\mu(C_0).$$

Therefore for each $n \in \mathbb{N}$ there is an l between $-N$ and N such that

$$\mu(C_l \cap C_n) > \tfrac{1}{2}\mu(C_0)\left(\frac{1}{2N+1}\right).$$

Take $\varepsilon = \frac{1}{2}\mu(C_0)/(2N+1)$. For each l between $-N$ and N, define

$$J_l = \{n \in \mathbb{N}: \mu(C_l \cap C_n) > \varepsilon\}.$$

Then some J_l has positive upper density, since $\bigcup J_l = \mathbb{N}$. □

Lemma 3.8. *Assume that $(E, \mathscr{E}, \mu)$, M, A and τ satisfy assumptions* (i)–(iv) *of Section* 2 *and that T and $\hat{T}$ are induced by τ as described in Section* 2.

If there is an $\varepsilon > 0$ and $J \subset \mathbb{N}$ such that

$$\mu(A \cap \tau^n A) > \varepsilon \quad \forall n \in J,$$

then there are $c \in \mathbb{R}$, $\delta > 0$ such that the inequality

$$P(M(A) \leq c, M(\tau^n A) \leq c) - P(M(A) \leq c)^2 \geq \delta \tag{7}$$

either holds for all but finitely many $n \in J$, or holds for all n in a set of positive upper density.

Proof. First suppose there is an atom D contained in infinitely many $\tau^n A$'s, $n \in \mathbb{N}$. Then by Lemma 3.2, there is a nonzero k such that $\tau^k D = D$. Now by Lemma 3.3 with

$$U = M(A \backslash \tau^n A), \quad V = M(\tau^n A \backslash A), \quad W = M(D), \quad R = M(A \cap \tau^n A \backslash D),$$

Inequality (7) holds for every $n \in \mathbb{N}$ that is a multiple of k.

Now suppose each atom is contained in only finitely many $\tau^n A$'s. Since the measures of the atoms in A must add up to a finite number, for all but finitely many $n \in J$ the continuous part of $A \cap \tau^n A$ has measure at least $\frac{1}{2}\varepsilon$. Therefore there exists an $F_n \subset A \cap \tau^n A$ with $\mu(F_n) = \frac{1}{2}\varepsilon$. Apply Lemma 3.3 with

$$U = M(A \backslash \tau^n A), \quad V = M(\tau^n A \backslash A), \quad W = M(F_n), \quad R = M(A \cap \tau^n A \backslash F_n).$$ □

4. Examples

In this section we construct examples of weakly mixing and mixing automorphisms using the 'stacking method'. First, however, we give an example relating to Theorem 3.6.

Example 4.1. *There exists a nonergodic automorphism τ for which T is a K-automorphism.*

Let $E=\mathbb{Z}\times\{1,2\}$, $\mathscr{E}=2^E$, let μ be counting measure, and let τ be the shift transformation $\tau\{(x,y)\}=\{(x+1,y)\}$. Let

$$A=\{(0,1),(1,1),(0,2)\}.$$

To see that A generates $(E,\mathscr{E})$, note that $\{(0,2)\}=A\backslash(\tau A\cup\tau^{-1}A)$ and $\{(0,1)\}=A\backslash(\{0,2)\}\cup\tau A)$. Thus the hypothesis of Theorem 3.6 is satisfied for any random measure which satisfies assumptions (i)-(iii). (Actually, assumption (iii) is not needed to prove that T is a K-automorphism.) But $\mathbb{Z}\times\{1\}$ is τ-invariant (with infinite measure), so τ is not ergodic.

The next two examples use the 'stacking' or 'interval-exchange' method of constructing automorphisms. We describe below how a transformation is constructed recursively using 'stacks' of subintervals of $[0,\infty)$. We will call τ an *infinite rank one* automorphism (by analogy with the finite case) because there is one stack of intervals at each stage. For a more rigorous description of this approach in the finite-measure case, see for instance Friedman (1970); the only difference between our construction of τ described below and the classical cutting and stacking construction is that in our case the measure is infinite.

We will take $(E,\mathscr{E},\mu)$ to be the half-line $[0,\infty)$ with Lebesgue measure on the Borel sets. We will define stacks of subintervals recusively — at the kth stage, we will have a stack C_k of height h_k:

$$C_k=(C_k(1),C_k(2),\ldots,C_k(h_k)),$$

where the $C_k(i)$'s are subintervals from $[0,\infty)$ of equal width, which we picture as stacked one above another. We write $\tilde{C}_k=\bigcup_{i=1}^{h_k}C_k(i)$. In our examples, we will take

$$C_1=(A)=([0,1)).$$

The stack C_k is constructed from C_{k-1} as follows: cut C_{k-1} into a given number s_k of subcolumns, each having the same width w_k. On top of each subcolumn, stack a finite number of disjoint intervals from $[0,\infty)\backslash\tilde{C}_k$ (where the new intervals have the same width as the subcolumns). Let $v(k,l)$ denote the number of intervals stacked on top of the lth subcolumn in order to construct C_k. These intervals should be chosen consecutively from $[0,\infty)\backslash\tilde{C}_k$, so that no part of $[0,\infty)$ is 'skipped'. Finally, stack each subcolumn on top of the one to the left. Thus each stack C_k consists of disjoint intervals of the same width.

The transformation τ_k is defined on $C_k(i)$, $i=1,\ldots,h_k-1$, by mapping each interval linearly to the one above it. Clearly each τ_k is an extension of τ_{k-1}; since intervals are chosen to have equal width, each τ_k is measure-preserving. If $\bigcup C_k=[0,\infty)$, then $\tau=\lim\tau_k$ is a well-defined automorphism on $(E,\mathscr{E},\mu)$.

Let $A=[0,1)$, $C_1=(A)$. For $k\geq 1$, define $\eta_k\in\{0,1\}^{\mathbb{Z}}$ by

$$\eta_k(i)=\begin{cases}1 & \text{if } 1\leq i\leq h_k \text{ and } C_k(i)\subset A,\\ 0 & \text{otherwise.}\end{cases}$$

If a vector η has only finitely many 1's, define $a(\eta)$ to be the position of the right-most 1 in η minus the position of the left-most 1 in η, plus 1. Roughly speaking, $a(\eta)$ is the 'height' of η disregarding leading and trailing zeroes. Define

$$\eta_k\cdot\eta_m=\sum_{i\in\mathbb{Z}}\eta_k(i)\eta_m(i).$$

Then $\eta_k\cdot\eta_m$ is the number of positions at which C_k and C_m each have a subinterval of A. Define the shift S by $S\eta(\cdot)=\eta(\cdot-1)$.

Define $\eta_k^{(l)}$ to correspond to the part of C_k from the lth subcolumn of C_{k-1}. More precisely, for $l=1,\ldots,s_k$,

$$\eta_k^{(l)}=S^m\eta_{k-1},$$

where

$$m=(l-1)h_{k-1}+\sum_{j=1}^{l-1}v(k,j).$$

Remarks. For any $k\geq 1$, the stack C_k is determined by η_k and $v(k,s_k)$. Given the parameters η_{k-1} and $v(k-1,s_k)$, the parameters s_k and $v(k,l)$, $l=1,\ldots,s_k$ determine the kth stack. One can see that if

$$\sum_{k=2}^{\infty}w_k\sum_{l=1}^{s_k}v(k,l)=\infty,$$

then τ is defined on $[0,\infty)$.

Also, if $v(k,s_k)$ is greater than $a(\eta_k)$, then τ^nA 'stays in' the stack as long as $0\leq n<a(\eta_k)$; more precisely,

$$\tau^nA\subset\bigcup_{i=n+1}^{h_k}C_k(i),\quad 1\leq n<a(\eta_k).$$

Hence

$$C_k(1)\cap\tau^nA=\emptyset,\quad 1\leq n<a(\eta_k).$$

It follows that if $C_1(1)=A$ and $v(k,s_k)$ is greater than $a(\eta_k)$ for each k, then

$$C_k(1)=A\Big\backslash\bigcup_{m=1}^{a(\eta_k)-1}\tau^mA,\quad k\geq 2.$$

This implies that if $C_1(1)=A$ then

$$C_k(1)\in\sigma\{\tau^mA\colon m\geq 0\}\quad\forall k\geq 1\tag{8}$$

and $\sigma\{\tau^nA\colon n\in\mathbb{Z}\}$ is the Borel σ-field on $[0,\infty)$.

Example 4.2. *There exists an infinite rank one weakly mixing transformation which is not mixing.*

Let $s_k = 2$ for all k; that is, the stack is cut in half at each stage. Take

$$v(k, 1) = 0, \qquad v(k, 2) > 2h_{k-1}.$$

By the remarks above, this choice of A and τ satisfies the hypothesis of Theorem 3.6 with E the positive half-line and $\mathscr{E}$ the Borel σ-field. We claim that τ is weakly mixing but not mixing.

In fact, τ is ergodic. As we stated in Section 3, on an infinite measure space ergodicity implies weak mixing. We use the following characterization of ergodicity (Friedman, 1970): τ is ergodic if and only if, for any $B_1, B_2 \in \mathscr{E}$ with positive measure, $\mu(B_1 \cap \tau^m B_2) > 0$ for some integer m.

Let x_1 and x_2 be Lebesgue points of B_1 and B_2, respectively. Then there is a $\delta > 0$ such that if J_1 and J_2 are intervals containing x_1 and x_2 whose lengths are less than δ, then

$$\mu(B_i \cap J_i) > \tfrac{1}{2}\mu(J_i), \quad i = 1, 2.$$

We can choose k large so that there are intervals in the kth stack having this property.

But $\tau^m J_1 = J_2$ for some integer m, by definition of τ. Hence $\tau^m B_1 \cap J_2$ and $B_2 \cap J_2$ each have measure strictly greater than $\frac{1}{2}\mu(J_2)$. Therefore $\mu(B_1 \cap \tau^m B_2) > 0$ and τ is ergodic.

To see that τ is not mixing, it is easy to verify (by looking at the kth stage of the stack) that for $n = h_{k-1}$, $k \geq 2$,

$$\mu(A \cap \tau^n A) = \tfrac{1}{2}.$$

We claim that τ and A satisfy assumption (iv) of Section 3. Obviously $\bigcup \tau^n A = [0, \infty)$. By the remarks preceding this example, $C_k(1)$ is in $\sigma\{\tau^n A \colon n \geq 0\}$ for all $k \geq 1$, and therefore A generates the Borel σ-field on $[0, \infty)$.

Therefore, by Theorem 3.6, if M is any random measure satisfying assumptions (i)–(iii) of Section 2, then T is weakly mixing but not mixing, and the same is true for $\hat{T}$.

Example 4.3. *There exists an infinite rank one mixing transformation.*

Let $A = [0, 1)$, $s_k = k$ (so $w_k = 1/k!$), and let $v(k, l)$'s satisfy

$$v(k, l) > a(\eta_{k-1}) + a(\eta_k^{(1)} + \eta_k^{(2)} + \cdots + \eta_k^{(l)}), \quad l = 1, 2, \ldots, s_k.$$

By the remarks preceding Example 4.2, this choice of A and τ satisfies the hypothesis of Theorem 3.6 with E the positive half-line and $\mathscr{E}$ the Borel σ-field. We claim that $\mu(A \cap \tau^n A) \to 0$.

For any $n > 1$, let $k = k(n)$ be such that

$$a(\eta_{k-1}) < n \leq a(\eta_k).$$

We will show that

$$w_k(S^n\eta_k\cdot\eta_k)\to 0 \quad \text{as } n\to\infty$$

and that this implies $\mu(A\cap\tau^nA)\to 0$.

Suppose $S^n\eta_k\cdot\eta_k>0$ for some n, $a(\eta_{k-1})<n\le a(\eta_k)$. Then there exist p, q in $\{1,2,\dots,k\}$ such that

$$S^n\eta_k^{(p)}\cdot\eta_k^{(q)}>0.$$

We claim there is only one such pair (p,q).

Observe first that since n is greater than $a(\eta_{k-1})$, p cannot equal q and hence p is strictly less than q.

Consider $S^n\eta_k^{(l)}\cdot\eta_k$ for $l=1,\dots,p-1$. Since $S^n\eta_k^{(p)}\cdot\eta_k^{(q)}$ is nonzero and

$$v(k,q-1)>a(\eta_k^{(1)}+\eta_k^{(2)}+\cdots+\eta_k^{(p)}),$$

the support of $S^n\eta_k^{(l)}$ lies to the right of the support of $\eta_k^{(q-1)}$ for each such l. But $v(k,p-1)$ is greater than $a(\eta_{k-1})$, so the support of $S^n\eta_k^{(l)}$ also lies to the left of the support of $\eta_k^{(q)}$. Therefore

$$S^n\eta_k^{(l)}\cdot\eta_k=0,\quad l=1,\dots,p-1.$$

Now consider $S^n\eta_k^{(l)}\cdot\eta_k$ for $l=p+1,\dots,q$. Since $S^n\eta_k^{(p)}\cdot\eta_k^{(q)}$ is nonzero and $v(k,p)$ is greater than $a(\eta_{k-1})$, the support of $S^n\eta_k^{(l)}$ lies to the right of $\eta_k^{(q)}$ for these values of l. But

$$v(k,q)>a(\eta_k^{(1)}+\eta_k^{(2)}+\cdots+\eta_k^{(q)}),$$

so the support of $S^n\eta_k^{(l)}$ lies to the left of $\eta_k^{(q+1)}$ for these values of l, if $q+1\le s_k$. Thus

$$S^n\eta_k^{(l)}\cdot\eta_k=0,\quad l=p+1,\dots,q.$$

Next consider $S^n\eta_k^{(l)}\cdot\eta_k$, for $l=q+1,\dots,s_k$. Since $S^n\eta_k^{(p)}\cdot\eta_k^{(q)}$ is nonzero and for each such l,

$$v(k,l)>a(\eta_k^{(1)}+\eta_k^{(2)}+\cdots+\eta_k^{(l)})+a(\eta_{k-1}),$$

the support of each such $S^n\eta_k^{(l)}$ lies to the left of the support of $\eta_k^{(l+1)}$ (if $l+1\le s_k$). Since n is greater than $a(\eta_{k-1})$, the support of $S^n\eta_k^{(l)}$ lies to the left of the support of $\eta_k^{(l)}$. Therefore

$$S^n\eta_k^{(l)}\cdot\eta_k=0,\quad l=q+1,\dots,s_k.$$

We have shown that

$$S^n\eta_k\cdot\eta_k=S^n\eta_k^{(p)}\cdot\eta_k\quad.$$

But $S^n\eta_k^{(p)}\cdot\eta_k^{(q)}>0$ and each $v(k,l)$ is greater than $a(\eta_{k-1})$, so $S^n\eta_k^{(p)}\cdot\eta_k^{(l)}$ is zero except when l equals q. Therefore

$$S^n\eta_k\cdot\eta_k=S^n\eta_k^{(p)}\cdot\eta_k^{(q)}.$$

It is easy to verify that

$$S^n \eta_k^{(p)} \cdot \eta_k^{(q)} \leqslant \eta_{k-1} \cdot \eta_{k-1} = (k-1)!,$$

so

$$w_k(S^n \eta_k \cdot \eta_k) = \frac{1}{k!}(S^n \eta_k \cdot \eta_k) \leqslant \frac{1}{k} \to 0 \quad \text{as } n \to \infty$$

(remember that k is a function of n).

Now $v(k, s_k)$ is greater than $a(\eta_k)$, so when the kth stack is shifted by an amount $n \leqslant a(\eta_k)$, it does not 'wrap around' the top; i.e.,

$$\mu(A \cap \tau^n A) = w_k(S^n \eta_k \cdot \eta_k).$$

Therefore $\mu(A \cap \tau^n A)$ converges to zero and, by Theorem 3.6, τ is mixing and so are T and $\hat{T}$ for any random measure M on $(E, \mathscr{E}, \mu)$ which satisfies assumptions (i)–(iii) of Section 2. (Actually, assumption (iii) is not needed to prove mixing.)

Remark. Recall that a sufficient condition for T to be a K-automorphism was that $\limsup \tau^n A$ have measure zero. This condition is not satisfied by Example 4.3; in fact, the tail σ-field $\mathscr{E}^\infty$ is all of $\mathscr{E}$. To see this, apply τ^n to both sides of (8); then $C_k(n)$ is in $\mathscr{E}^n$. For $k \geqslant 2$, let $n = n(k) = h_{k-1} + 1$ and define

$$B_n = \bigcup_{j=n}^{h_k} (C_k(k) \cap A).$$

Now observe that B_n is in $\mathscr{E}^n$ since for each j, $(C_k(j) \cap A)$ is either $C_k(j)$ or empty. But

$$\mu(A \cap B_n) = (k-1)(k-1)!/(k!) \to 1 \quad \text{as } k \to \infty.$$

Therefore $A \in \mathscr{E}^n$ for every n, and $\mathscr{E}^\infty = \mathscr{E}$. In particular, $\mu(\limsup \tau^n A) \neq 0$.

References

P. Billingsley, Probability and Measure (Wiley, New York, 1986).

S. Cambanis, C.D. Hardin, Jr. and A. Weron, Ergodic properties of stationary stable processes, Stochastic Process. Appl. 24 (1987) 1–18.

S. Cambanis, A. Ławniczak, K. Podgórski and A. Weron, Ergodicity and mixing of symmetric infinitely divisible processes, Tech. Rep. No. 346, Center for Stochastic Processes, Dept. of Statist., Univ. of North Carolina (Chapel Hill, NC, 1991).

I.P. Cornfeld, S.V. Fomin and Ya.G. Sinai, Ergodic Theory (Springer, New York, 1982).

N.A. Friedman, Introduction to Ergodic Theory (Van Nostrand Reinhold, New York, 1970).

U. Krengel and L. Sucheston, On mixing in infinite measure spaces, Z. Wahrsch. Verw. Gebiete 13 (1969) 150–164.

G. Maruyama, The harmonic analysis of stationary stochastic processes, Mem. Fac. Sci. Kyushu Ser. Mat. IV 1 (1949) 49–106.

G. Maruyama, Infinitely divisible processes, Probab. Theory Appl. 15 (1970) 3–23.

B.S. Rajput and J. Rosinski, Spectral representations of infinitely divisible processes, Probab. Theory Rel. Fields 82 (1989) 451–487.

P. Walters, An Introduction to Ergodic Theory (Springer, New York, 1982).